Vol. XII, No. 4
November 1961

CALIFORNIA JOURNAL OF EDUCATIONAL RESEARCH

IN THIS ISSUE:

published by
CALIFORNIA TEACHERS ASSOCIATION
1705 Murchison Drive Burlingame California

CALIFORNIA JOURNAL OF EDUCATIONAL RESEARCH

Published by the CALIFORNIA TEACHERS ASSOCIATION for the
CALIFORNIA ADVISORY COUNCIL ON EDUCATIONAL RESEARCH

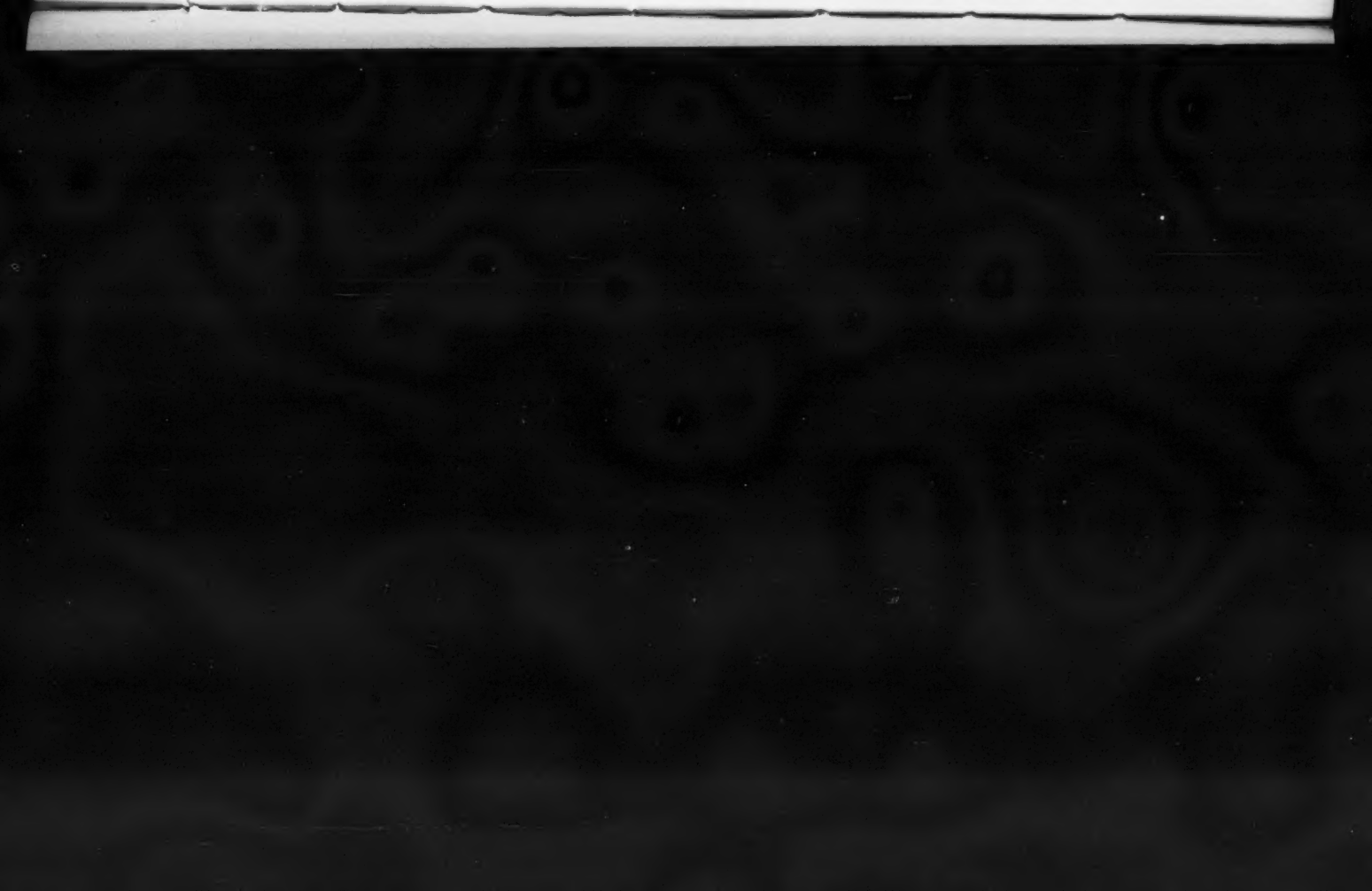
●

Published at Burlingame five times a year: January, March, May, September, and November. Subscription price: $6.00 per year; single copies, $1.50. All correspondence should be directed to: CALIFORNIA JOURNAL OF EDUCATIONAL RESEARCH, 1705 Murchison Drive, Burlingame, California. Checks should be made payable to the CALIFORNIA TEACHERS ASSOCIATION.

●

Contents listed in Education Index
Certain articles listed in Psychological Abstracts
Member of the Educational Press Association of America

Entered as second class matter at the Post Office at Burlingame, California



CALIFORNIA JOURNAL OF EDUCATIONAL RESEARCH

Volume XII SEPTEMBER 1961 Number 4

CONTENTS

THE EDITORS SAY:

Philosophy and Research

J. R. Shannon

Philosophy and research are two radically different approaches to truth in education, and to the dilettante they may appear to be wholly separate and unrelated. But the mature scholar knows better: philosophy and research, although different, are closely related and each dependent on the other. Research depends on philosophy for guidance, and philosophy depends on research for substance.

Research first depends on philosophy for guidance in its planning stage. All research is predicated on certain basic assumptions, and all evaluation in research is in terms of certain criteria, but both the assumptions and the criteria are arrived at by philosophical thinking. There is an obvious example of this relationship between philosophy and research in education in our concern about the efficiency of the high school.

Are high schools offering (and requiring) enough courses in science, mathematics, and foreign languages? Do the courses have the proper content? Are pupils mastering the subjects as they should? The words, **enough, proper,** and **should** throw these questions into the province of philosophy as much as into the province of research. The most simple survey procedures will tell us what and how many courses in the respective areas are offered and required, the contents of the courses, and how well the pupils are mastering the courses, but research can never tell us whether the courses are **enough,** whether their contents are **proper,** and whether pupils' mastery of them is what it **should** be. That is the function of philosophy.

By considering and adopting the basic assumptions and deciding the criteria, philosophy lays the track on which research runs. In fact, some professional philosophers claim that fixing assumptions and criteria is their chief function. But, just as philosophy precedes research by laying the track, so does it follow by deciding the implications and applications of research findings.

A good illustration of research's need for philosophical guidance in the implications and applications of its findings was provided by a professor in

(Continued on page 192)

Dr. Shannon, Emeritus Professor, Indiana State Teachers College, Terre Haute, Indiana, is well known through his extensive writing in the field of education. After serving as teacher, principal, superintendent, director of research, and college teacher in the mid-west, Dr. Shannon became the Coordinator of Graduate Studies at Sacramento State College (1948-52).

Learning Abilities in Mexican-American And Anglo-American Children

ARTHUR R. JENSEN

The apparently poor scholastic aptitude and achievement of a large proportion of the children from certain ethnic or national groups in the United States is a long recognized problem which seems to gain in importance over the years. The problem is especially conspicuous to public school personnel in parts of California and the Southwest with respect to Mexican-American children (10). Recently it was noted by the investigator that in a number of California school districts many of the Mexican-American children are being classified as "slow learners" or as "mentally retarded" on the basis of currently popular standard intelligence tests, and are often placed in special classes.

In the Mexican-American group the "slow learners" are not the exception, but are a large majority. A variety of intelligence tests, both verbal and non-verbal, generally classify the majority of the Mexican-Americans considerably below the average of the predominantly Anglo-American[1] normative groups on which the tests were standardized. For example, in a number of studies the average Stanford-Binet IQ's of Mexican-Americans tend to cluster around 80 (1, 2, 4, 6, 8, 11). It is noteworthy that the use of a Spanish translation of the Stanford-Binet with the bilingual Mexican-Americans does not change the picture essentially (9). The apparent intellectual or scholastic inferiority of these children lies deeper than the language handicap of not knowing the words used in the verbal parts of

[1]The term *Anglo-American* is used throughout this paper to refer to white, American-born children from monolingual, English-speaking families.

***Arthur R. Jensen has been assistant professor of educational psychology at the University of California, Berkeley, for three years. His previous experience includes one year as clinical psychologist at the University of Maryland Psychiatric Institute and two years as a USPHS postdoctoral research fellow at Maudsley Hospital, University of London. Dr. Jensen, who obtained his Ph.D. degree in 1956 from Teachers College, Columbia University, is a co-author of* From Adolescent to Adult.**

most intelligence tests. Also, it is doubtful that the main clue to the problem is in the fact that the majority of the Mexican-American children come from families of lower socio-economic status, against which many standard intelligence tests tend to discriminate to some extent (3). In some of the schools investigated, many of the Mexican-American and Anglo-American children come from families that are indistinguishable with respect to social class as determined by the usual indices, such as father's occupation, neighborhood of residence, parents' education, and income. But there are still very marked IQ and scholastic differences between the two groups.

Teachers of the mentally retarded have at times noted that on the playground the Mexican-American children appear to be normally bright as compared with the Anglo-American retarded children, even though all are educationally retarded. Thus one immediately suspects a basic difference underlying the intellectual handicap of the two groups.

It was with these observations in mind that the investigator conceived of the method used in the present study for assessing the educational potentialities of Mexican-American children, as well as those of other national subcultures and ethnic groups, by using tests which provide direct measures of present learning ability. The standard intelligence tests currently in use are actually static measures of achievement which sample the knowledge and skills the child has acquired in the past. Often this sampling can be quite inappropriate for children who have not had much exposure to the Anglo-American culture of the normative group. It would seem that a better way to measure learning potential, or to decide whether a child is inherently a slow learner, would be to give the child a standard task and observe how fast he learns it, whereby we may note how readily the child's behavior changes through the trial-and-error of experience. Aside from reliability, there are two chief requisites of such a "test" if it is to be useful for the purpose we have in mind: (1) that language facility in itself is not to be a crucial variable in determining performance; (2) that the materials of the learning task and the nature of the task itself be equally familiar and comprehensible to children from a variety of subcultures.

Some simple learning tasks which seem to fulfill these requirements were devised. The question with which the present study was concerned was whether or not groups of Mexican-American and Anglo-American children equated on IQ, as measured by the *California Test of Mental Maturity*, are also equal in learning ability, as measured by certain standard learning tasks. The study was carried out under the author's supervision by three graduate students in educational psychology. While these experiments can be considered as a pilot study, they do permit some important conclusions to be made before large-scale studies are conducted in this area. It would seem especially important that the diagnostic potentiality of further developments of the kinds of tests used in this study be brought to the attention of other workers in this field.

Learning Tasks

The learning materials were of two types, referred to as *Familiar* and *Abstract*. There were three types of learning tasks: *Immediate Recall, Serial Learning,* and *Paired-Associates Learning.*

Familiar Materials. There were two equivalent forms of the test using the familiar materials, which consisted of 12 common, readily recognized objects. The objects in Form A were: water glass, a bar of soap, a paintbrush, a candle, a ball, a wristwatch, a plastic toy airplane, a comb, a toy gun, a toy car, a spoon, and a rubber dagger. Form B consisted of: a doll, a fork, a miniature chair, a pair of sun glasses, a small plastic horse, a pair of scissors, a toy boat, a plastic cowboy, a cup, a key, a book and a doll's shoe.

Abstract Materials. These materials consisted of seven plastic forms approximately 1¼ inches in diameter and ½ inch thick and of the following shapes and colors: yellow diamond, green diamond, yellow triangle, blue triangle, red circle, blue circle, and green square.

Recall Test. This test always preceded the *Serial* or the *Paired-Associates* tasks. Its main purpose was to familiarize the subject with the materials. The experimenter sat opposite the subject at a low table. In a carton box, the inside of which was never visible to the subject, was contained the set of objects of either the *Abstract* or the *Familiar* tests. Each object was placed one at a time directly in front of the subject, who was asked to name the object. It did not matter whether he named the object correctly, as long as it was named. If the subject could not name it, the experimenter supplied the name. (In the present study this was never necessary in the *Familiar* tests and was necessary in only a few cases with the *Abstract* test.) The subject could name the objects in Spanish if he wished. After the subject had named all the objects in the set, he was told to look them over carefully because they were to be put back in the box to see how many could be recalled. The subject was given 10 seconds to study the objects; they were then put back into the box and the subject was asked to name as many as he could recall. As each object was recalled, the experimenter took it from the box and placed it before the subject. When he could not recall any more objects for 30 seconds, the remainder were placed before him one at a time and he was required to name each as he had done in the preliminary trial. The entire procedure was repeated until the subject could recall every one of the objects. Subject's score was the total number of unrecalled items before attaining criterion.

Serial Learning. After the *Recall* test, the experimenter placed a set of 12 (7 in the Abstract test) inverted plain white cardboard boxes (4" x 4" x 8") in a row before the subject, who was turned around so he could not

see the objects placed under each of the boxes. The subject then had to learn which object was under each box, by starting at his left end of the series, guessing what was under each box in turn, and lifting the box to see if he was correct. The subject went through the series repeatedly in this manner until he could name the object under every box, never being permitted to depart from the serial order of left to right. Subject's score was the total number of errors made before attaining criterion. The serial order of the objects was different for each subject within each experimental group, but the same orders were used in each group.

Paired-Associates Learning. This test made use of Forms A and B of *Familiar* objects. One set of 12 objects was fastened to the outside of each cardboard box, with the other set placed under the boxes. Thus there were 12 boxes with 12 pairs of objects—one object in each of the boxes, and one object fastened to the outside of each. The subject's task was to learn what was inside each of the boxes. His only clue was the object on top of the box, since the order of the boxes was completely rearranged on each trial to rule out serial learning. The *Paired-Associates* task was preceded by the *Recall* test for the objects placed inside the boxes. Subject's score was the total number of errors made before attaining the criterion of correctly guessing the objects under all 12 boxes.

Subjects

All subjects were selected from the fourth and sixth grades in five public schools in Contra Costa County, California.[2] These schools draw mainly from a semi-rural, lower socio-economic laboring-class population, a large proportion of which is Mexican-American. The subjects in this study were selected on the basis of several factors: language spoken in the home, intelligence as measured by the *California Test of Mental Maturity*, age, and socio-economic status of the family.

All subjects were native-born Americans and were from the lower socio-economic class, as judged from the father's occupation, locality of residence, and the teacher's knowledge of the child's home. Their fathers were mostly unskilled and semi-skilled laborers. The Mexican-American group was bilingual, with Spanish spoken in the home. The Anglo-American group came from monolingual English-speaking families. The study used only children from homes in which both parents were living, and children with any kind of physical disability were not selected.

It should be mentioned that although the Anglo-American subjects were

[2]We are indebted to Mr. Francis G. Burke, the Director of Educational Services in Contra Costa County, and to Miss Zelma Parker, the School Psychologist, for making available the subjects and facilities for this study.

selected at random within the conditions called for by the experimental design, it was very difficult to find enough Mexican-Americans to fill one of the cells in the experimental design, even though the subjects were being drawn from every fourth-grade class in four large schools with large proportions of Mexican-Americans (approximately one-fourth to one-third). In order to fill the cell requiring nine Mexican-Americans with IQ's over 110, it was necessary to obtain four of the subjects from a large school in a neighboring county. It would be virtually impossible to conduct a large-scale study of this type if the experimental design called for matching Mexican-Americans with Anglo-Americans along the entire range of IQ's: In any one school one could find only a few Mexican-Americans with above average IQ's.

Experiments

I. The purpose of this experiment was to compare the learning scores on both the Familiar and Abstract tasks for *Recall* and *Serial Learning* in groups of "Bright" and "Dull" Mexican-Americans and Anglo-Americans. The Dull group had *California Test of Mental Maturity* IQ's ranging from 73 to 89; the Bright group had IQ's in the range of 111 to 126. Table I shows the mean IQ's and ages in each group. All subjects were selected from the fourth grade and ranged in age from 9 to 11.

TABLE I

Mean Age and IQ of Sample Groups

Group	*N*	*Mean Age*	*Mean IQ*	*SD*
Bright Mexican-American	9	9.33	117.33	6.10
Bright Anglo-American	9	9.11	115.67	4.27
Dull Mexican-American	9	9.67	82.89	5.82
Dull Anglo-American	9	9.78	81.78	3.93

Half the subjects were given the *Recall* and *Serial* tests, with both *Familiar* and *Abstract* objects, by Mrs. Arlene Rustin Gundzik, and half were tested by Mrs. Joanne Clason. Half the subjects took the *Abstract* tests first and half took the *Familiar* tests first. It was found that the order of administering the tests had no significant effect. The tests were administered during a single session, which took, on the average, approximately one hour.

II. In this experiment the Mexican-American and Anglo-American groups were selected so as to span the IQ range from 60 to 120 or above.

Because of the scarcity of high IQ's among the Mexican-Americans in the fourth grade (they were all used in Experiment I), subjects for this experiment were selected from the sixth grade. This grade was preferable to either the third or the fifth, since more recent IQ tests (CTMA) were available. (The children are routinely tested in the 2nd, 4th, and 6th grades.) The same conditions for selection mentioned earlier were applied. There were two Mexican-Americans and two Anglo-Americans in each of four IQ levels: 60-70, 80-90, 100-110, and 120 or above. Thus there was a total of 16 subjects in Experiment II. Miss Jacqueline Rapier tested them on the *Recall, Serial Learning,* and *Paired- Associates* tasks, using only the *Familiar* objects.

III. The purpose of this experiment was to determine the equivalent forms reliability of the *Recall* and *Serial* tasks using the *Familiar* objects, the tests which showed the most promise in the earlier experiments. Miss Rapier administered to every child in a single fourth-grade class (N = 22) both the *Recall* and *Serial* tests, using Forms A and B two weeks apart. The class was typical for the school district; six of the twenty-two pupils were Mexican-Americans, all considerably below the class average in IQ.

Results

Experiment I

The performances of the four groups can be shown most simply in graphical form. Everything represented in Figures 1 to 4 was tested for statistical significance by analysis of variance. The analysis revealed no significant main effect for "Nationality." However, there was a significant ($p < .05$) *"Nationality"* × *IQ* interaction, with the Dull Mexican and Dull Anglo-American groups widely separated in learning scores, and the Bright groups of both Mexican-Americans and Anglo-Americans more similar. The low IQ Mexican-Americans did about as well as both of the Bright groups. The results are shown graphically in Figures 1 to 4. It was thought that perhaps there would be a *"Nationality"* × *Test* interaction, with the Abstract tests tending to favor the Anglo-American group, but this interaction turned out non-significant in the analysis of variance. Only the *Recall* tests showed a significant over-all difference ($p < .05$) between IQ levels. Also the significance of the *"Nationality"* × *IQ* interaction was largely attributable to the *Recall* tests.

The correlation between *Recall* and *Serial Learning,* based on the composite Abstract and Familiar scores of all subjects, was .64 ($p < .01$). The correlations between the learning of Abstract and Familiar objects (based on the composite *Recall* and *Serial* scores) was .68 ($p < .01$). These correlations would indicate rather substantial reliability of these tests.

Experiment II

The *Recall, Serial,* and *Paired-Associates* tests differed widely in difficulty. *Recall* was easiest and *Paired-Associates* was by far the most difficult. The mean number of errors on each test for all subjects was: *Recall,* 7.5; *Serial Learning,* 26.1; *Paired-Associates Learning,* 73.5. This finding is interesting in itself. Why should the *Paired-Associates Learning* be so much more difficult than *Serial Learning*? In both cases twelve "facts" or S-R connections had to be learned. Probably the greater structure or organization of the *Serial* task made it easier to learn than the *Paired-Associates.*

Analysis of variance performed for each of the tests separately showed that, with the small number used in this experiment, only the *Serial Learning* test achieved statistically significant differences. The other tests were consistent with the *Serial* test, as would be expected with an intra-class correlation between the tests of .75, but the differences found with the *Recall* and *Paired-Associates* tests fell short of significance at the .05 level. The *Serial* test, on the other hand, showed significant differences between IQ levels, and showed a significant *"Nationality"* × *IQ* interaction, with the low IQ Mexican-Americans performing much better than the low IQ Anglo-Americans, while the two groups showed little difference at the high IQ levels. In short, the picture was essentially the same as that found in Experiment I and shown graphically in Figures 1-4.

Experiment III

Since the *Serial Learning* test seemed to evince the most promise for further research and development, it was decided to determine its reliability by means of an equivalent form administered after a two-week's interval. Since the *Recall* test always preceded the *Serial* task in the previous experiments, it was used in this study also, and its reliability was determined.

Reliability was computed by the Hoyt method, using analysis of variance (Guilford [7], pp. 383-385). For the *Serial Learning* of Familiar objects, the reliability of composite scores based on the two forms of the test was .88; the reliability of scores on a single form was .81. These reliability coefficients compare favorably with those of standard intelligence tests. The composite score reliability of the *Recall* test, on the other hand, was only .42; and for a single form it was only .26. It appears that the nature of the objects had a considerable effect on the ease with which they can be learned for *Recall.* But once the *Recall* task had been mastered, the nature of the objects had little effect in the *Serial Learning.* The reliability of *Recall* scores for any one form of the test is probably higher than the reliability obtained from "equivalent" forms would indicate. The fact is, of course, that we did not actually have equivalent forms with respect to *Recall* learning. Since the *Recall* scores seem to be more a function of the test materials (i.e., one would predict a considerable *materials* × *subjects*

Figure 1

Mean error scores of Dull (IQ's 73-89) and Bright (IQ's 111-126) Mexican-American and Anglo-American fourth grade children on the Recall test for Familiar objects.

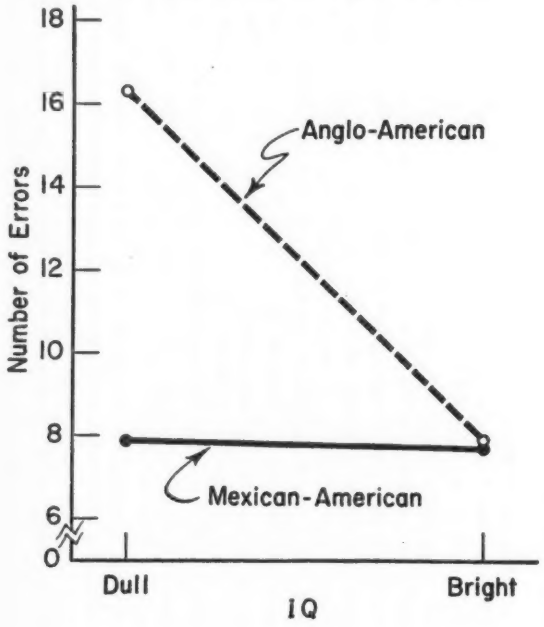


Figure 2

Mean error scores of Dull and Bright Mexican-American and Anglo-American children on the Serial Learning of Familiar objects.

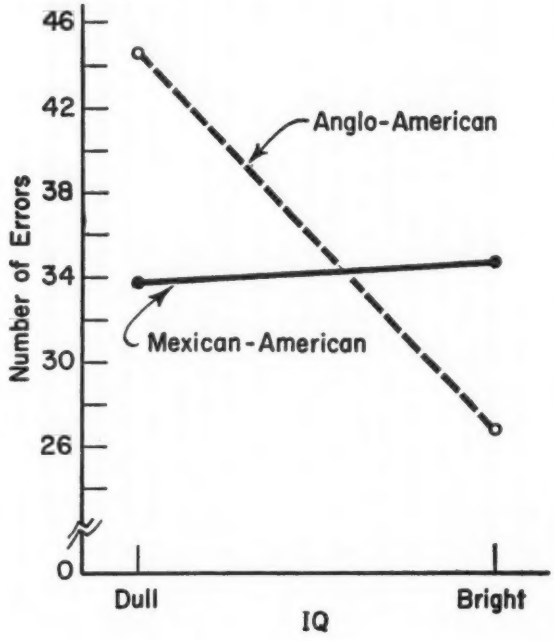


Figure 3

Mean error scores of Dull and Bright Mexican-American and Anglo-American children on the Recall test for Abstract objects.

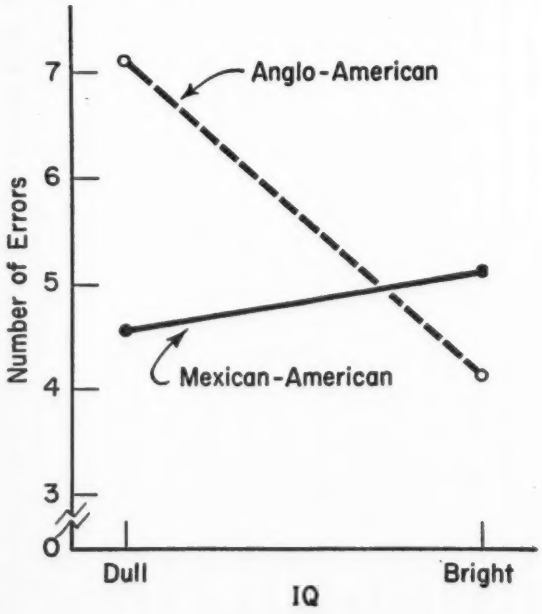
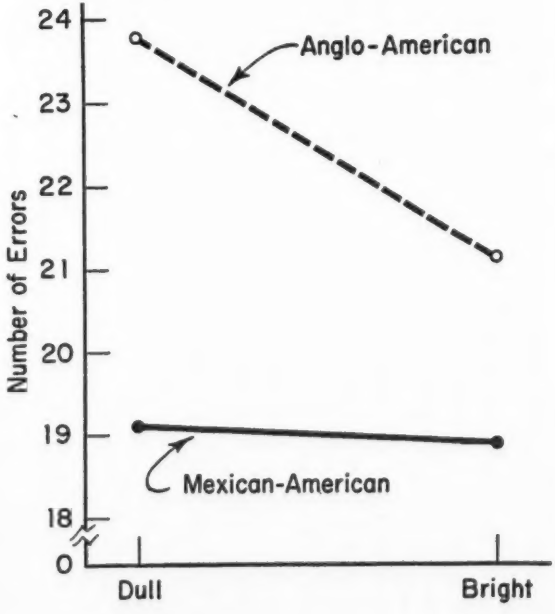


Figure 4

Mean error scores of Dull and Bright Mexican-American and Anglo-American children on the Serial Learning of Abstract objects.



interaction) one would place more confidence in the *Serial Learning* task as a more generalized and stable measure of learning ability . The reason for the relatively poor showing of the *Paired-Associates* test in Experiment II is obscure. It was noted by the experimenters that the *Paired-Associates* test took much longer than any other test and there seemed to be a waning motivation and an increasing frustration in a number of the subjects during the *Paired-Associates* testing. These tendencies were not observed in the *Recall* and *Serial Learning*. A shorter form (i.e., fewer items) of the *Paired-Associates* task might have produced better results. This should be tried in a future study.

The correlation between IQ (CTMA) and *Serial Learning* scores in this class is only —.30, which is non-significant for this small sample. (The correlation is negative because the learning score is expressed as number of *errors*.) Since both the IQ and *Serial Learning* tests have high reliability, this degree of correlation between the two indicates that the *Serial Learning* task is not just another IQ test, but is a relatively independent measure of learning ability.

Discussion of Findings

The most important finding of these experiments is that on the particular learning tasks used in this study, Mexican-American children with low IQ's performed significantly better than Anglo-American children with low IQ's. In fact, low IQ Mexican-Americans performed as well on the learning tests as did those of high IQ in both the Mexican-American and Anglo-American groups. The IQ test (*California Test of Mental Maturity*) seems to have been successful in discriminating between fast and slow learners in the Anglo-American group, but did not do so in the Mexican-American group. The fact that high IQ Mexican-Americans are so rare in the population from which our samples were drawn, along with the fact that they performed no differently on the learning tests than did the low Mexican-Americans, suggests that the IQ test is discriminating in the Mexican-American group on some other basis than that on which it discriminates in the Anglo-American group. The IQ in the Anglo-American group is a valid index of learning potential; the low IQ's among the Anglo-Americans were indeed the slow learners. But this is not necessarily so in the Mexican-American group. Considering the very large proportion of low IQ's in the Mexican-American population from which this study drew its samples, the findings of the study are consistent with the hypothesis that the distribution of basic learning abilities in Mexican-Americans is not substantially different from that in the Anglo-American population of comparable socio-economic level. Apparently the IQ tests did not help us to find, in the total fourth-

grade population, any really slow learners among the Mexican-Americans. This fact should not be surprising if one considers the small size of the sample and the fact that the true slow learners constitute only 5 to 10 per cent of the population. In the Anglo-American population, on the other hand, we were quite able to identify the slow learners by means of the IQ test.

The question then arises, why do the low IQ Mexicans, if they are not all really slow learners, appear to their teachers to be retarded in their school work? And why is the IQ able to predict their poor scholastic performances? The hypothesis that first comes to mind is that the low IQ Mexican-American children have not acquired in their environment the kinds of knowledge, habits, and skills that provide the basis for school learning and which are tapped by IQ tests. We have evidence that an elaborate matrix of previous learning underlies a person's verbal learning ability. Frequency of past exposure to verbal stimuli is correlated with ease of learning where verbal materials are involved. For example, it has been shown that the speed of learning a list of nonsense syllables is highly correlated with the frequencies of the letters in our language that compose the nonsense syllables (12). Non-verbal intelligence tests probably discriminate unfairly against the Mexican-Americans almost as much as do verbal tests for the reason that most of the non-verbal performance tasks require verbal mediation. Reasoning, problem solving, and concept formation depend largely upon verbal behavior, whether or not the actual problems involved are "verbal" in the sense that they consist of words. A so-called non-verbal test, such as *Raven's Progressive Matrices*, for example, actually requires a great deal of verbal mediation on the part of the subject. No one who has ever taken or administered "non-verbal" tests can fail to note the great amount of subvocal, and often overtly vocal, behavior made in response to test items.

If the IQ tests are used to discriminate the slow learners with respect to school learning so that they can be placed in special classes for the retarded, it becomes important in groups such as the Mexican-Americans to make a more refined diagnosis to determine whether they are slow learners in the same sense that the low IQ Anglo-Americans are slow learners, or whether they are slow learners in particular school subjects only because of some cultural handicap. Certainly the two types of slow learners must need quite different kinds of educational treatment. On the learning tests used in this study, for example, we found that Mexican-Americans with IQ's in the sixties are very fast learners as compared with Anglo-Americans in this IQ range. Yet both groups are classified as "retarded" and are subjected to the same teaching methods. The present study would suggest that this is clearly a mistake. We cannot say as yet just what the nature of the remedial instruction should be for the Mexican-Americans, but it seems safe to say that after a differential diagnosis has been made with direct tests of learning ability, the low IQ but normal-learning Mexican-Americans should be treated differently than the retarded Anglo-Americans.

With direct learning tests we would undoubtedly also find a certain small proportion of low IQ Anglo-Americans who are not truly slow learners. Their educational needs are probably still different.

The present study indicates that simple learning tasks may be developed for this type of differential diagnosis of educational disabilities. One of these tests—*Serial Learning*—had a reliability as high as that of many standard intelligence tests. But the tests do not measure IQ, as indicated by the low correlation found between IQ scores and learning scores. The question of validity does not seem a problem at present, for these tests are patently measures of learning. The main question that remains to be answered by further research is whether other kinds of learning tests might prove more effective. What we can most likely expect is that a number of different types of learning tasks will be found that correlate only to a moderate degree with each other, but which, when taken together, will correlate highly with some external criterion. Garrett, for example, found that eight different learning tests, each of which had a low correlation with the other tests and with IQ, when taken together had a multiple correlation with IQ of .60 (5). Thus, it is likely that a battery of learning tests will have to be developed in order to achieve the kind of diagnostic instrument we would desire for use in schools with children of ethnic and cultural backgrounds for whom the usual intelligence tests are of limited value.

Summary

Groups of Mexican-American and Anglo-American fourth and sixth grade school children of different IQ levels ranging from 60 to 120 or above were compared on a number of learning tasks consisting of *Immediate Recall, Serial Learning,* and *Paired-Associates* learning of Familiar and Abstract objects. The tests are not proposed for individual diagnostic use at present, since more research is needed for a more complete understanding of their value, and at present there is not sufficient normative data to provide a basis for interpreting individual scores.

The main finding is that on the direct measures of learning ability used in this study, Anglo-American children of low IQ are slow learners as compared with Mexican-Americans of the same IQ. Mexican-Americans of above average IQ do not differ significantly in learning ability from Anglo-Americans of the same IQ. However, high IQ's are very rare among the Mexican-American population. This study suggests that the majority of Mexican-Americans with low IQ's, at least as measured by the *California Test of Mental Maturity,* are actually quite normal in basic learning ability, though they may be poor in scholastic performance for reasons other than inherently poor learning ability. A low IQ in the Anglo-American group, on the other hand, is in most cases a valid indication of poor learning ability.

It was suggested that most of the low IQ Mexican-Americans, not being basically slow learners, should not be placed, as they now are, in classes with the Anglo-Americans of low IQ, who are basically slow learners and therefore require different methods of teaching.

At least one of the tests used in this study (*Serial Learning*) had an equivalent forms reliability as high as that of most standard IQ tests. The development of a complete battery of direct learning tests would seem to have considerable promise for improving the diagnosis of educational disabilities, especially in ethnic and cultural groups for whom the usual intelligence tests are not highly appropriate.

REFERENCES

1. Coers, W. C. "Comparative Achievement of White and Mexican Junior High School Pupils," *Peabody Journal of Education*, 12: 157-62, 1935.
2. Cook, J. C. and Arthur, G. "Intelligence Ratings for 97 Mexican Children in Saint Paul, Minnesota," *High School Journal*, 38:24-31, 1955.
3. Eells, K. *Intelligence and Cultural Differences.* Chicago: University of Chicago Press, 1951.
4. Garretson, O. K. "A Study of Causes of Retardation Among Mexican Children in a Small Public School System in Arizona," *Journal of Educational Psychology*, 19:31-40, 1928.
5. Garrett, H. E. "The Relation of Tests of Memory and Learning to Each Other and to General Intelligence in a Highly Selected Adult Group," *Journal of Educational Psychology*, 19:601-13, 1928.
6. Garth, T. R. and Johnson, H. D. "The Intelligence and Achievement of Mexican Children in the U. S.," *Journal of Abnormal Social Psychology*, 29:222-39, 1934.
7. Guilford, J. P. *Psychometric Methods.* New York: McGraw-Hill, 1954.
8. Hill, H. S. "Correlation Between IQ's of Bilinguals of Different Ages on Different Intelligence Tests," *School and Society*, 44:89-90, 1936.
9. Keston, M. J. and Jiminez, C. A. "A Study of the Performance on English and Spanish Editions of the Stanford-Binet Intelligence Test by Spanish-American Children," *Journal of Genetic Psychology*, 85:263-69, 1954.
10. Record, W. "Racial Diversity in California Public Schools," *Journal of Negro Education*, 28:15-25, 1959.
11. Sheldon, W. "The Intelligence of Mexican Children," *School and Society*, 19:139-42, 1929.
12. Underwood, B. J. and Schultz, R.W. *Meaningfulness and Verbal Learning.* Chicago: Lippincott, 1960.

Individual Differences and Effectiveness Of Auto-Instruction at The Primary Grade Level

JOHN D. MCNEIL and EVAN R. KEISLAR

Recent studies (1,2,4) have suggested that achievement and retention may not correlate with individual differences when learners are taught by self-instructional materials. These studies point up the need for examining the classical predictors of successful learning in connection with programs of auto-instruction. Further, the relationship between individual differences and liking of auto-instruction invites investigation.

The major hypothesis of the present study was that certain individual differences are significantly related to effectiveness of auto-instruction. Effectiveness was operationally defined as (a) achievement on a post-test, and (b) expressed preference for auto-instruction following the learning opportunity. In addition, attention was given to the relationship between individual differences and four techniques for presenting the auto-instructional program.

Subjects and Procedure

The study was conducted in two parts. In the first part, 123 children in the first and second grades were given a 13-day auto-instructional course in principles of molecular theory. The children were divided into two matched groups. Those in one group were individually taught by an auto-instructional device which required an overt response; that is, they had to respond

Evan R. Keislar is professor of education at the University of California, Los Angeles. Before coming to this institution in 1949, Dr. Keislar served as assistant professor at Tufts College, instructor at Princeton University, and research associate with the College Entrance Examination Board. He obtained his Ph.D. degree in 1944 from the University of California, Berkeley.

John D. McNeil, associate professor of education at UCLA for the past four and a half years, also serves as associate director of student teaching. Formerly a research assistant at Columbia University Institute of Field Studies and a teacher in the San Diego City Schools, Dr. McNeil obtained his Ed.D. degree in 1956 from Columbia University.

correctly to each item presented by pressing one of three buttons in order for the program to advance. Those in the second group individually received the same program but were not required to respond overtly; they simply observed the program as it was presented by the device.

The independent variables used in this study were (a) scores on the Kuhlmann-Anderson Intelligence Test, (b) teacher's assessment of the child's motivation to learn (i.e., attentiveness, persistency in task, and the asking of relevant questions), and (c) a readiness test indicating the child's familiarity with the illustrations to be used as contacts between everyday events and the theoretical principles to be taught. The dependent variables were an opinionnaire which assessed children's attitudes toward the program and a post-test consisting of objective questions and problems. Both were administered immediately following the course.

The second part of the study, carried out in another school, was essentially a replication of Part One with these additions:

1. Children in Grade 3 were substituted for second graders.

2. A matched uninstructed control group was established. Members of this group did not receive instruction in molecular theory but were taught aspects of mathematics in an audio-visual presentation which included the assignment of problems involving manipulation of blocks.

3. Two additional experimental groups were formed and given the auto-instructional program by a technique whereby the content was projected simultaneously to an entire group by slides and accompanying taped commentary. This latter arrangement also permitted confirmation of a correct answer or statement through exposure of the answer together with the stimulus panel. One of the groups received the programmed sequence in a question form; the other in a statement form. Content and sequence were identical for those in the collective situations and those individually operating machines.

The 216 pupils who served as subjects during Part Two were matched on the basis of IQ, sex, and grade level before assignment to either the control group or one of the four experimental situations. In addition, available information concerning parental occupations (scientific-technological vs. social service) was subsequently collected.

Results

1. **Achievement.** As seen in Table I, mental age is very clearly related to final achievement scores based upon a test which called for the application of principles of molecular theory to new situations. The correlations which range from .89 to .43 are all significant.

TABLE I

Correlation Between Mental Age and Achievement Following Auto-Instruction of Primary Grade Children in Two Schools

SCHOOL A	*Variables*	*Mean*	*Standard Deviation*	*Number*	*r*
Grade 1	Mental Age	6.3	1.04	60	.89
	Achievement	42.03	9.243		
Grade 2	Mental Age	7.7	.693	63	.65
	Achievement	49.9	10.669		

SCHOOL B	*Variables*	*Mean*	*Standard Deviation*	*Number*	*r*
Grade 1	Mental Age	6.8	1.19	70	.45
	Achievement	42.9	7.937		
Grade 3	Mental Age	8.6	1.83	89	.43
	Achievement	50.1	7.071		

A child's motivation to learn was determined from a five-point rating scale completed by his teacher. Correlations between motivation and achievement for the pupils in the four grades described in Table I were as follows:

School A Grade 1 $r = .69$ Grade 2 $r = .62$
School B Grade 1 $r = .45$ Grade 3 $r = .29$

The test of readiness administered only in School A showed correlations with achievement of .60 in Grade I and .60 in Grade 2.

A difference was obtained between the achievement scores of boys and girls; boy learners in each of the four groups who had been matched with girls on the basis of mental age did significantly better on the post-test ($t = 4.35$, $P < .01$). This is a noteworthy finding in view of the fact that boys and girls did not differ in their initial readiness for the program or knowledge of the content to be taught as revealed by (a) the readiness test given in School A and (b) the post-test scores made by the control group at School B.

Occupation of parents was not shown to be significantly related to achievement. Neither was the technique used for presentation of the program found to have a differential achievement effect for children with particular characteristics.

2. **Attitudes toward the auto-instructional program.** Children's attitudes toward the program were shown by their responses to such questions as: *"Do you want to go on learning by machine?" "Which would you prefer to do—(a) stay in your room and make something, (b) listen to your teacher read a story, or (c) learn some more things by machine?"* Only the teacher's assessment of the child's motivation to learn was associated with expressed desires to continue the program; 89 per cent of the more highly motivated (those with ratings in top two categories on the teacher rating scale) and 72 per cent of the less motivated children indicated that they wanted to continue the auto-instructional program. This difference was significant at the .05 level ($X^2 = 3.94$). Although there was no significant difference between those with high and low motivation with respect to their preference for auto-instruction rather than the alternate activity of hearing a story or making something, 53 per cent of those with low motivation chose auto-instruction as compared with 39 per cent of those with high motivation.

Preference for the topic taught ("molecules") rather than the other topics dealing with either concrete or abstract subject matter ("trees" or "numbers") was positively associated with mental age. Eighty-four per cent of the children in the top one-third of the population in terms of mental age selected "molecules" as a subject for further study, while only 64 per cent of the bottom one-third chose this subject ($X^2 = 5.17$, $P < .05$).

Conclusions

1. When taught molecular theory by an auto-instructional program, individuals of lower and higher mental ability are distinguishable in terms of a criterion performance calling for transfer of principles even though nearly all children display evidence of having profited from instruction.

2. There may be two reasons why classical predictors of successful learning apply to this auto-instructional program but conflict with conclusions reached in the recent studies mentioned previously. First, unlike those programs which taught subject-matter at a recall level, e.g., spelling, the criterion performance in this study required the development of a generalization and its application in new situations. This criterion is clearly related to mental age. Secondly, the mean number of errors made during this program may have been higher than the error rates in the other studies. Previous analysis of the characteristics of the molecular course, reported elsewhere (3), revealed an average error rate of 14 per cent for children in the first grade.

3. The discovery that sex is a relevant factor in achievement, measured by the criterion measures of this study, points up the importance of considering sex as a variable in inquiries regarding the learning of specific

kinds of subject matter. The fact that boys did better than girls is of particular interest since it cannot be argued that under auto-instruction boys received a preferential treatment in connection with a scientific topic as might be the case under ordinary classroom conditions. Because the use of an auto-instructional device guarantees impartiality to all pupils, it appears to be a desirable instrument for research in the study of personal characteristics and learning.

4. There is no evidence that the individual differences identified in this study are related to effectiveness of varied auto-instructional techniques. The method of program presentation, for example, showed no interaction with mental age.

BIBLIOGRAPHY

1. Feister, Charles B. and Stanley M. Sapon, "An Application of Recent Developments in Psychology to the Teaching of German," *Harvard Educational Review*, Volume 28, No. 1, Winter 1958.

2. Irion, A.L. and L. J. Briggs, "Learning Task and Mode of Operation Variables in Use of the Subject Matter Trainer." Lowry Air Force Base, Colorado. Air Force Personnel and Training Research Center, October 1957. Technical report AF PTR C-TR 57-8 (abstracted in *Teaching Machines and Programmed Learning—a Source Book*, edited by A. A. Lumsdaine and Robert Glaser, National Education Association, 1960, p. 62.

3. Keislar, Evan R. and John D. McNeil. "Teaching Scientific Theory to First Graders by an Auto-Instructional Device," *Harvard Educational Review*, Volume 31, No. 1, Winter 1961.

4. Porter, Douglas, "Some Effects of Year-Long Teaching Machine Instruction," *Automatic Teaching: The State of the Art* (edited by E. T. Galanter), New York: John Wiley & Sons, Incorporated, Chapter 7, 1959, pp. 85-90.

Differing Views of Institutional Aims Among College Administrators, Teachers, and Students

JOSEPH P. RICE, JR.

A statement of purpose containing the aims and philosophy of an institution is considered basic, and such a statement usually is displayed prominently in a school's publications. Most institutions devote considerable effort to revision and implementation of their philosophy through committee work and reappraisal. In general, the stated philosophy of a junior college tends to be stylized and similar to those of other higher institutions, except that more emphasis is placed on terminal vocational programs and adult education.

It is recognized that a given institution's philosophy is the end result of many traditional, empirical, and idealistic influences and represents a summary point of view. As such, personal disagreements are to be expected. However, even casual discussion with administrative, faculty, and student groups reveals wide variation of opinion concerning the aims a college should serve. Both within-group and cross-group differences can be noted. Each group appears to have a unique philosophical outlook and anticipates different results from a college education.

The purpose of this study was to discover how divergent are the philosophical outlooks of junior college administrators, teachers, and students. Some of the questions which prompted the study were: *Do administrators, teachers, and students hold mutually exclusive opinions concerning the aims of higher education? How disparate are the cross-group differences? Can certain qualitative generalities be made concerning these differences?*

Method of Study

Utilizing various sources including: (1) junior college catalogues, (2) lists of purposes proposed by experts, and (3) literature pertaining to college

Joseph P. Rice is psychologist for the Lompoc Unified School District, a position he has held for a year and a half. Formerly testing officer at Yuba College, Marysville, and instructor at Fresno City College, Dr. Rice obtained his Ph.D. degree in 1959 from the University of Connecticut.

philosophy, a check list of possible purposes for the existence of a college was composed. Lamar Johnson's "twelve goals" (6), Jesse Bogue's "objectives" (2), and the "eleven purposes of general education" of the President's Commission on Higher Education (11), were especially useful sources. From these and other sources, a final list of 17 possible purposes for a college's existence was compiled. These are shown in Table I and are preceded by the short titles used in succeeding tables. With the exception of those relative to guidance, the statements were considered mutually exclusive and qualitatively different. Five guidance aims were included in order to explore this area more minutely. Each of the 17 possible purposes were rated as *"major," "minor,"* or *"not"* purposes for a college's existence by groups of junior college administrators, faculty members, and students. Also, subjects were asked to rank the purposes from: (1) *"most essential"* to (17) *"least essential"* for a college's existence. A crude measure of relia bility was obtained by comparing the coincidence of highly rated purposes with highly ranked ones.

Initially, 15 junior colleges were selected for study. The schools were of differing size, location, and population in order to be representative of the universe of California junior colleges. The sample of administrators was drawn from all 15 schools because of the relatively small number in this group, while the teacher and student samples were drawn from three of the fifteen junior colleges, and included a large, middle-sized, and small school. A sampling of teacher and student groups from a wide variety of junior colleges was deemed unnecessary, since significant variation within given groups from different schools was not found (using the chi square test). In fact, significant variation was not found between sub-groups of students (terminal vs. transfer) or between sub-groups of teachers (academic subjects vs. vocational subjects) using the chi square test. Likewise, significant variation was not found to be attributable to school size or location.

The final samples included: (1) The deans or presidents of 15 junior colleges, (2) thirty junior college teachers from three schools (ten from each), and including twenty teachers of academic subjects and ten teachers of terminal subjects, and (3) sixty sophomore junior college students, including 38 transfer and 22 terminal students. The samples were considered to be fairly representative of California junior colleges generally. The teachers, chosen at random by an alphabetical selection process, were given the rating sheets with an explanation and returned the sheets directly by mail to the writer. The rating sheet was administered to students in their psychology or sociology classes by the writer.

An earlier study (12) utilizing this check list compared differences between private and public junior college administrators. In this study, undertaken as part of a broader study of over-all college guidance programs, it was found that private and public junior college administrators varied considerably in their ratings and rankings of college aims, although within

TABLE I

Seventeen Possible Aims of a College

Brief Form	*Description*
Academic Guidance	Provides counseling and guidance for students with academic problems.
Adult Education	Provides a broad program of adult education including an evening division beyond high school.
Aesthetic	Promotes the discovery, development, and use of aesthetic skills and encourages appreciation and enjoyment of aesthetic outlets such as music and art.
Basic Skills	Teaches and places emphasis upon those subjects which include the basic skills necessary for everyday living.
Creativity	Emphasizes the discovery and exploitation of creative talents and encourages unique expression including the appreciation of uniqueness in others.
Democratic Life	Promotes the democratic way of life emphasizing the study and perpetuation of democratic institutions.
Formal Education	Provides preparation for further formal education and background.
General Guidance	Provides general orientation and guidance programs for all students.
Homemaking	Prepares young people for homemaking and family life.
Intellectual	Emphasizes the use of the intellect and promotes critical thinking.
Mental Health	Promotes self-understanding and good mental health habits.
Peace	Promotes international understanding and peace.
Personal Guidance	Provides help for the student in solving personal adjustment problems.
Social Maturity	Promotes social development and adustment to everyday society, helping the student to gain poise, self-assurance, and mature social bearing.
Spiritual	Helps the student discover and develop sound moral and spiritual values.
Vocational Education	Provides specific vocational education preparing the student for employment in specific skilled areas.
Vocational Guidance	Provides counseling and guidance based on an objective assessment of a student's ability, interests, and needs in order to help him choose his life's work.

each group consistency was reliable. In general, the ratings of administrators of private colleges resembled the ratings of the teacher group in the present study.

During another study of students in the writer's classes, it was found that subjects tended to rate the purposes differently depending upon whether they understood the ratings to represent an appraisal of their "actual schools" or the "ideal school." Although the results might have been

clearer had the question been put more specifically, most subjects seemed to have understood the inquiry to concern their actual school's functioning.

Discussion of Findings

A summary of the findings is contained in Tables II and III which show how the samples rated and ranked college aims.

The significance of the percentages reported in Table II depends, of course, on the size of the samples. Therefore, comparisons between the teachers and students have more significance than between administrators and others. Applying the test for significance of difference between proportions, most of the differences of 8 per cent or better may be considered significant. As a further test for significant differences, the chi square test was applied to each datum after grouping the "*minor*" and "*not*" categories together and using an all-group means as the "expected" figure. Those

TABLE II

Three Group Ratings of 17 Aims as *Major, Minor* or *Not* Purposes for Their Colleges

PURPOSE	*PERCENT OF SUBJECTS RATING PURPOSES AS:*[1]								
	Major			*Minor*			*Not*		
	A.[a]	*T.*[b]	*S.*[c]	*A.*[a]	*T.*[b]	*S.*[c]	*A.*[a]	*T.*[b]	*S.*[c]
Academic Guidance	93*	70	60	7	30	32	—	—	8
Adult Education	100*	33*	78*	—	60	18	—	7	4
Aesthetic	40*	60*	12*	47	27	64	13	13	24
Basic Skills	93*	43	35	7	37	49	—	20	16
Creativity	33*	67*	17*	53	30	61	13	3	22
Democratic Life	100*	90	35*	—	10	44	—	—	21
Formal Education	100	93	88	—	7	10	—	—	2
General Guidance	93*	47	51	7	33	47	—	20	2
Intellectual	100	93	32*	—	7	54	—	—	14
Homemaking	53*	40*	26*	40	33	50	7	27	24
Mental Health	47*	86*	21*	53	7	69	—	7	10
Peace	7*	53*	15*	80	27	48	13	20	37
Personal Guidance	80*	67	58	20	20	27	—	13	15
Social Maturity	93*	53*	17*	7	27	50	—	20	33
Spiritual	80	80	15*	20	10	38	—	10	47
Vocational Education	100*	33*	86*	—	33	12	—	33	2
Vocational Guidance	100*	47	55	—	33	40	—	20	5

[a] Junior College Administrators, N = 15

[b] Junior College Teachers, N = 30

[c] Junior College Students, N = 60

[1] To nearest percent

*Chi square significant when compared with both other groups.

percentages followed by an asterisk in Table II differ significantly at the .05 level or better from both other groups applying the chi square test.

Since the mean ranks in Table III can be treated as means, it is possible to establish significance by using the t-test for significance of difference between means. Figures which differ significantly from both of the other groups are designated. Most of the discussion of "differences" which follows is based on the significant differences explained above.

In general, the most marked differences in ratings appear between students and others. The students tended to be more critical and used the *"minor"* and *"not"* categories with far more frequency than did the others. Whereas the administrators and the teachers tended to view all of the aims as being *major,* or at least *minor,* aims of a college, the students came very close to rejecting a few as simply not being aims of a college. This may indicate a tendency for administrators and teachers to recite abstract philosophy while failing to implement their aims in working theory and practice. Since so many students rejected some aims, it may be that they have received no introduction to them as part of their programs. It must be

TABLE III

Ranking of 17 College Aims from (1) Most Essential to (17) Least Essential

Purpose	*Administrators N = 15*			*Teachers N = 30*			*Students N = 60*		
	Mean Rank		*S.D.*	*Mean Rank*		*S.D.*	*Mean Rank*		*S.D.*
Academic Guidance	10[a]	(9.9)*[b]	3.4	7	(7.8)*	3.1	5	(5.8)*	3.1
Adult Education	6	(8.2)*	2.8	9	(9.2)*	4.7	3	(2.8)†	1.3
Aesthetic	15	(13.8)	2.2	11	(10.6)*	2.4	15	(14.6)	1.3
Basic Skills	4	(5.6)†	1.0	10	(9.8)	3.7	8	(8.7)	2.8
Creativity	16	(14.6)†	.9	6	(7.1)†	2.7	13	(12.0)†	1.7
Democratic Life	8	(8.9)*	2.4	3	(3.2)†	.6	10	(10.1)*	3.5
Formal Education	2	(1.9)	.4	2	(2.5)	.5	1	(1.3)	.4
General Guidance	7	(8.7)	3.0	12	(11.2)*	2.9	7	(8.1)	3.2
Homemaking	12	(10.9)	3.1	17	(15.8)†	1.1	12	(11.4)	2.2
Intellectual	9	(9.4)	2.5	1	(1.7)†	.6	9	(9.5)	2.6
Mental Health	13	(11.2)	1.8	4	(4.3)†	1.2	11	(10.9)	1.8
Peace	17	(15.9)	1.1	13	(12.3)*	2.1	16	(15.7)	.8
Personal Guidance	14	(11.7)*	3.4	8	(8.5)	3.3	6	(6.9)	2.9
Social Maturity	5	(7.9)†	3.6	14	(13.2)	1.8	14	(13.4)	1.6
Spiritual	11	(10.8)†	1.0	5	(6.4)†	2.8	17	(16.2)†	.4
Vocational Education	1	(1.4)	.5	16	(15.0)†	.8	2	(1.8)	.7
Vocational Guidance	3	(3.1)	.2	15	(14.4)†	1.2	4	(3.7)	1.5

[a] Composite rank or ranks of the mean ranks

[b] Applying the t-test, mean rank differs significantly (*) or very significantly (†) from both of the other groups.

remembered that all the students were sophomores with ample exposure to their school.

In general, the ranking method proved to be a sensitive way to distinguish among the groups studied. More agreement was found between administrators and students (rho = .70) than between administrators and teachers (rho = —.14) or teachers and students (rho = .09). Since these correlations were based on composite figures, they are of dubious meaning and significance. The method of forced ranking is, of course, open to serious criticism, since the ranker may actually accept or reject all of the propositions put to him but yet must treat them as if there were a "best" and "worst." This criticism was somewhat offset, however, by having the subject first rate the qualities to be ranked. Hence, it can be seen that since administrators tended to rate most aims as *"major"* and rarely as *"not,"* their rankings must be interpreted within a different framework than those of the students who far more frequently rated aims as *"not."* The fact remains, however, that all of the groups were able to pick the three or four most essential aims with fair ease, just as they were able to pick the few least essential aims. It was only in the middle ranks that any difficulty of choice was observed.

The administrative viewpoint: Administrators tended to rate and rank most highly the more practical and widely accepted college aims. This is understandable, since they are instrumental in the writing of catalogue descriptions and other information concerning college aims. None of the aims listed came close to being rejected by administrative ratings. Their rankings disagreed most with those of the teachers. General, broad, and practical aims such as "vocational education," "formal education," and "vocational guidance" were ranked high. Abstract, personal, or theoretical aims such as "personal guidance," "aesthetic," and "creativity" were ranked lowest. In general, the administrators agreed more with one another than did teachers or students.

The teachers' viewpoint: The teachers tended to rate all of the aims highly, although they were somewhat more critical than were the administrators. Even though they accepted most of the aims as being *major,* they disagreed considerably with the other groups over which were most essential. Teachers ranked the more impractical and abstract aims such as "intellectual," "democratic life," and "mental health" as most essential (an exception was "formal education" which was also ranked among the most essential), while ranking more practical aims as "general guidance," "vocational guidance," and "vocational education" as least essential. Many of the aims highly regarded by teachers, such as "creativity," "democratic life," "intellectual," "mental health," and "spiritual," were considered to be least essential by students.

The students' viewpoint: Administrators and teachers have been exposed

to many influences, such as meetings, education, and professional literature. These exposures would naturally affect their perception and recognition of the aims a typical college ought to have. No such indoctrination can be assumed for students, and they tend to view the aims from a framework of their own needs and expectations. The more practical and obvious aims such as "formal education" were rated and ranked highly. Many aims, such as "social maturity," "aesthetic," "peace," and "spiritual" were roundly rejected.

Interestingly, the students appeared to serve notice that they wanted no interference with their value systems. This would seem to put them at odds with teachers who, as a rule, indicated the essentiality of dealing in the area of values. The only differences between the rankings of administrators and students were in the areas of "academic and personal guidance" and "social maturity"; the students ranking "guidance" aims as more essential and the administrators ranking "social maturity" as more essential.

Summary of Findings

Groups of college administrators, teachers, and students rated and ranked 17 college aims in terms of their importance and essentiality. Obvious and fundamental differences among the groups were noted. Teachers were found to deviate more markedly from the others in terms of how they ranked the aims for essentiality. The teachers tended to rank the more abstract aims more highly than did the others. Students tended to be more critical of the aims generally, while administrators tended to be least critical.

The administrators rated all aims highly, while ranking the more practical aims as being more essential. Teachers also rated most of the aims highly but ranked them in a different order than did the administrators or students. The more visionary aims were deemed most essential by teachers, while the more practical aims were put on a lesser level of essentiality. Students, like the administrators, ranked the practical aims most essential, but, in addition, rejected the abstract aims of the teachers more decisively than did the administrators.

The three groups were shown to have three distinguishable and different hierarchies of aim. One wonders if more inter-group communication and open discussion on the subject of aims isn't indicated. While most of the institutions studied have very similar, and even stereotyped, stated philosophies, the groups clearly differed in their acceptance of their school's aims.

The following secondary findings are of special note:

1. Students tended to rank all types of guidance as more essential than did the others.

2. Subsequent subjective interviews revealed that students rejected interferences by an institution in the spiritual, moral, psychosexual (home-

making), social and maturity areas.

3. All of the groups, but especially the students, rejected the peace aim.

4. The teachers seemed to stand alone in their positive emphasis on the intellectual, mental health, spiritual, and creative aims.

5. Administrators tended to agree with students on the most essential aims (formal education, vocational guidance) and the least essential (aesthetic, peace, creativity), but differed in degree of emphasis of the middle-ranked aims.

6. Teachers differed in most aspects with both groups.

It is recommended that further studies of college aims and differences of viewpoint be undertaken in more diverse settings such as universities and technical schools. Perhaps future research will want to create a more exhaustive list of aims.

BIBLIOGRAPHY

1. Bateman, Harold C. and Stone, D. R. "Basic Elements in Defining General Education," *Junior College Journal*, 27:90-92, October 1956.

2. Bogue, Jesse P. *The Community College*. New York: McGraw-Hill Book Co., 1950. 390 pages.

3. Bright, H. F., and Colvert, C. C. "Research Problems Preferred by Junior College Administrators," *Junior College Journal*, 20:350-54, February 1950.

4. Brownell, Samule. "Trends of Higher Education," *Journal of Education*, 137:3-22, February 1955.

5. Harvard Commission on Objectives of General Education in a Free Society. *General Education in a Free Society*. Cambridge: Harvard University Press, 1946. 267 pages.

6. Johnson, Lamar. *General Education in Action*. Washington: Council on Education, 1952. 604 pages.

7. Koos, Leonard V. "Junior College Administrators and Their Scope of Function," *School Review*, 111:143-50, March 1955.

8. Koos, Leonard V. "The Junior College's Full Scope of Function," *Junior College Journal*, 18:55-56, October 1947.

9. Lowenstein, Norman and Yates, Vivian. "How Freshmen Feel about College," *Personnel and Guidance Journal*, 33:379-80, March 1954.

10. Present and former Member of the Faculty. *The Idea and Practice of General Education*. Chicago: University of Chicago Press, 1950. 333 pages.

11. President's Commission on Higher Education. *Higher Education for American Democracy*. New York: Harper and Bros., 1947. 6 v.

12. Rice, Joseph P. *A Comparative Study of the Guidance Program in the Private and Public Junior College*. Storrs: University of Connecticut, 1959.

MEET YOUR CACER -- 9

(This is the ninth of a series of biographical sketches of the members of the California Advisory Council on Educational Research.)

Transplanted from Pennsylvania Dutch country at the age of 22 months (he claims credit for persuading his parents to "go West"), D. Welty Lefever has been a Southern Californian ever since, except for periods when he taught at out-of-state universities. Welty attended Pasadena public schools, La Verne College, and the University of Southern California, where, in 1927, he received the first Ph.D degree granted in any department by that institution. He can take credit for breaking ground in another area also, having been one of the original members of CACER at its inception in January 1948. He has been a loyal and productive member of the Council ever since.

After teaching at La Verne College and Compton High School, where he also served as counselor and registrar, Dr. Lefever joined the faculty of USC's School of Education and this spring completed his 35th year in that institution.

In addition to his duties as Professor of Education at USC, Welty is active in a variety of other professional endeavors. He is faculty associate with both the Youth Studies Center (the Ford Foundation Research center at USC) and the Anaheim instructional television experiment; is a member of the Board of Trustees of La Verne College; is on the Education Committee of the California Institution for Women at Corona; and is one of the trustees of the HEAR Foundation, Los Angeles.

Dr. Lefever's name can be found as the author or co-author of many professional works: monographs on the Thorndike Intelligence and Rorschach tests, an experiment on an interneship program for teachers, the SRA Achievement Series, and a general textbook on guidance.

Welty's favorite forms of relaxation are music—including his collection of folk song recordings—and detective stories in all media. Always a proud father of a son who is now a high school mathematics teacher and basketball coach, Welty is now an even prouder grandfather. An avid photographer, he is now combining his two favorite "hobbies" and producing a picture record of the life and times of one Eric Bruce Lefever.

The Predictive Value of the College Ability Test For A Group of College Freshmen

PHYLLIS E. KENNEDY

One of the functions of college testing and counseling departments is to evaluate the examinations currently in use in the college-wide testing program. The usefulness of the tests used for predicting students' successful continuance in college and for general advisement purposes should be determined.

The primary purpose of this study was to determine the relationship between the College Ability Test (SCAT) and the academic achievement of college freshmen as measured by their grade-point averages and final grades in selected freshman courses. In studies of this nature, freshmen grade averages have generally been used as the criterion of academic success in lieu of a more adequate measure. Since course grades and grade averages were among the more readily available measures, they were selected as the criteria in the present study.

Specifically, the following data were obtained: (1) The correlations of the College Ability Test Verbal, Quantitative, and Total scores with first-semester grade-point averages; and (2) The correlations of the College Ability Test scores and final grades in language arts, history, physical science, and psychology courses.

Procedure

Description of the measures. The SCAT was administered from May to September to all new freshmen applying for admission in the fall semester. The SCAT is a 70-minute multiple-choice test which attempts to measure "developed ability" in verbal and quantitative skills.

Four basic freshman courses were selected from the major areas of the required general education courses listed in the college bulletin: *History 255,* a course in "The United States in the Modern World;" *Physical Science 150,* facts and principles of chemistry, physics, and the scientific method;

Phyllis E. Kennedy is testing supervisor at San Fernando Valley State College, where she has served since 1956. Formerly a psychometrist at the Veterans Administration Guidance Center at Los Angeles State College, Miss Kennedy obtained her M.S. degree in education from the University of Southern California in 1955.

Language Arts 150, a basic course in communications skills; and *Psychology 150,* "Principles of Human Behavior." All were three-unit courses.

The criteria of achievement were the first-semester grade-point averages of the total group of new freshmen who were enrolled for ten or more units, and the final grades earned by freshmen enrolled in the selected courses. The grade averages were calculated on a four-point scale with a unit of "A" earning four grade points, a "B" three points, a "C" two points, a "D" one point, and a grade of "F" zero grade points.

Method. Two hundred ninety newly enrolled freshmen in the Fall 1959 semester at San Fernando Valley State College were studied. All test data cards for new entering freshmen were taken from the files. The total number of units attempted, the number of grade points earned, and the grade-point averages were computed and were entered on each test data card. Final grades in the selected courses were also recorded.

Correlations of SCAT scores with semester grade-point averages were computed for the 290 students who attempted ten or more units. A tally was prepared in order to compare the majors selected by the men and women.

Correlations of SCAT scores with final course grades were obtained by using the course grades of all freshmen in *History 255* (N=52) and in *Physical Science 150* (N=68), 103 freshmen in *Psychology 150,* and the first 100 freshmen with course grades in *Language Arts 150.* Freshmen in these groups were enrolled for at least seven units at registration.

The data were inspected and met the requirements for product-moment correlation.

Results of Study

Table I presents the correlations of the SCAT Verbal, Quantitative, and Total scores with the first-semester grade-point averages for men and women freshmen and for the combined group who were enrolled for ten or more units.

Correlation coefficients of .34, .18, and .35 for grade averages and SCAT Verbal, Quantitative, and Total scores, respectively, were obtained for the male students. The correlations of the grade averages with the Verbal and Total scores were significant.

Correlations between grades and all three test scores were significant for the group of women. Correlations of grades with Verbal, Quantitative, and Total scores were .45, .27, and .46, respectively. Verbal and Total scores predicted achievement about equally well. A clearer picture of the relationship between test scores and grades was obtained when separated by sex than when the groups were combined.

Differences in the size of the coefficients for men and women may be

TABLE I

The Correlations of Semester Grade Averages and College Ability Test Verbal, Quantitative, and Total Scores for College Freshmen

	Men (N=100)		*Women (N=190)*		*Total Group (N=290)*	
Test	*r*	*S.E.*	*r*	*S.E.*	*r*	*S.E.*
SCAT—Verbal part	.34*	.11	.45**	.08	.40**	.06
SCAT—Quant. part	.18	.12	.27**	.04	.17*	.07
SCAT—Total score	.35**	.11	.46**	.08	.37**	.07

*Significant at the .05 level.
**Significant beyond the .01 level.

TABLE II

Per Cent of Men and Women Freshmen with Majors in Each College Division

Major Division	*Per Cent of Men Freshmen*	*Per Cent of Women Freshmen*
Business and Economics	21	7
Education	5	46
Fine Arts	9	8
Health, P.E., Rec.	4	5
Languages and Literature	7	9
Science and Mathematics	34	9
Social Sciences	16	13
Major Unknown	4	3
Total	100%	100%

TABLE III

The Correlations Between Final Course Grades and College Ability Test Verbal, Quantitative, and Total Scores for College Freshmen

		College Ability Test Scores		
Course	*N*	*Verbal*	*Quantitative*	*Total*
Language Arts	100	.44**		
History	52	.31*	.18	.33*
Physical Science	68	.45**	.53**	.62**
Psychology	103	.43**	.40**	.45**

*Significant at the .05 level.
**Significant at the .01 level.

related to such factors as motivation, attitudes, interests, and employment. It is also possible that the selection of majors is related to the higher coefficients obtained for women students, even though freshmen do not specialize their first year. Table II indicates the major divisions the freshmen selected.

One-third of the men were majoring in the Science, Mathematics, and Engineering Division, one-fifth in Business and Economics, and 16 per cent in the Division of Social Sciences. Almost half of the women listed a major in the Education Division (primarily elementary education), with the next largest group (13 per cent) in Social Sciences.

SCAT were also found to be significantly related to final grades in the four selected courses. These results are given in Table III.

It will be observed that the highest validity coefficients were obtained for the physical science course, and the next highest for the psychology course. Somewhat lower correlations were found with the language arts and history courses.

When the three SCAT correlations were compared, the Total score yielded somewhat higher correlations with final course grades than Verbal or Quantitative, at least in the particular courses that were selected in this study. It is interesting to note that the Quantitative score was related to student performance in certain types of courses but lost discriminating ability when correlated with the semester grade averages of the total freshman group.

It appears that the SCAT scores are somewhat more useful for predicting final course grades than in predicting the semester grade-point averages. As some selections had occurred before the students were admitted, the group was more homogeneous than an unselected group, and this had a tendency to reduce the size of the validity coefficients. The range of ability was further limited by students who withdrew from college early in the semester for academic reasons. The lower relationship between the test measures and the semester grade-point averages than between test scores and final grades also suggests that factors such as variations in grading and the combining of many majors have tended to blur the findings.

REFERENCES

1. *Cooperative School and College Ability Tests Technical Report.* Princeton: Educational Testing Service.

2. *Examiner's Manual Cooperative School and College Ability Tests.* Princeton: Educational Testing Service, 1955.

3. Guilford, J. P. *Fundamental Statistics in Psychology and Education.* Second Ed., New York: McGraw-Hill, 1950.

The Use of the Q-Technique in Determining Curriculum Content

WILLIAM J. SCHILL

The curriculum content of the public schools is constantly being evaluated. Faculty committees, curriculum experts, and citizens' advisory committees are concerned with the courses offered and their content. Still other methods are used to decide upon curriculum content for courses preparing students for entry occupations. Industrial surveys to determine curriculum content are common practice in the United States and take many forms, covering areas from future manpower needs (2) to the skills necessary for industrial technicians (6). The primary instrument in the typical industry survey is a rating scale, requiring respondents to rate a variety of offerings in a particular curriculum.

If public education is to be improved in a systematic fashion, it is essential that the attempts to define curriculum content be as sophisticated as possible. A procedure which is getting more and more attention in educational evaluation and research is Stephenson's Q-technique (9), which has proven useful in applied psychology. The use of the Q-technique as a group measure was outlined by Morsh (7); Cohen (3) has provided details of simplified methods for computing the Q-technique correlations; and more recently a paper by Sorenson and Sheldon (8) presented the use of the Q-technique in classroom evaluation.

The purpose of this paper is to give an example of the use of the Q-technique in determining the curriculum content and to explain the development of the Q-sort, its use, and the analysis. The example will be followed by a discussion of the advantages in using the Q-technique.

Example

This study was undertaken to determine the mathematical concepts or skills used by electronics technicians in California industries. The purpose

---

William John Schill is a research consultant at the University of California, Los Angeles. His former experience includes that of teaching assistant (statistics), teacher and department chairman in technical education (electronics), and public high school teacher. Currently a doctoral candidate at UCLA, Mr. Schill obtained his M.A. degree in 1952 from the University of Minnesota.

was to define the needed mathematical competencies in order to determine what mathematics should be included in junior college curricula designed to train students for jobs as electronics technicians.

Constructing the Q-Sort

To identify the mathematical continuum, the author used a variety of studies that have already been published, including David Allen's study conducted in California (1). Allen's study listed approximately 100 mathematical skills or competencies needed in the electrical-electronics industries in California. In addition, standard texts in the area, final examinations administered at the junior colleges in the State, and interviews with 25 teachers of electronics who were attending the summer session at U.C.L.A. were used. Each of the teachers interviewed was asked to submit five problems which would characterize five mathematical concepts or skills that they would expect their graduates to be able to master. In addition, the author applied his knowledge of mathematics in electronics. All of the skills and concepts accumulated from the aforementioned sources were grouped into those that were similar or the same, arranged in a logical sequence, and inspected as to whether or not they were needed. Although addition, subtraction, multiplication, and division of whole numbers are needed in the electrical-electronics industries in California and are undoubtedly used continually by electronics technicians, it was assumed that these fundamental arithmetic skills are essential to everyone in our modern industrial society.

There evolved 66 mathematical concepts, ranging from the more complex arithmetic processes to integral and differential calculus. The next step in the process of building the Q-sort was to develop typical problems that represented these specific skills or concepts. The 66 problems were screened to be sure that the language of the electrical-electronics industry had been omitted, and that as often as possible the single mathematical skill or concept was represented by the problem.

Pilot Study

At this point, a Pilot Study was conducted, the function of which was to:

1. Check upon the time that it took the technicians to do the sort. (This proved to be between 45 and 70 minutes.)

2. Check on the validity of the problems to determine whether they truly represented the mathematical skill or concept that they were intended to represent.

Eighteen technicians in the electrical-electronics industries in Southern California were interviewed. They were selected technicians who had been identified as having mathematical skills and who would be, and proved to be, able to recognize the skill or concept involved in each problem. After the interview and after the sort, the author had an extensive discussion with each technician to see whether he could identify the specific mathematical skill or concept that each of the 66 problems represented. These interviews resulted in the changing of a number of problems. There were instances in which the author had unwittingly included electrical or electronic language in the symbols used in the problem. An example of this is the use of "db" as a literal number, whereas in electronics it represents decibels. Another example is the use of "S" as a literal number, whereas in electronics it represents the complex "j" factor. After each interview, the Q-sort problems were re-evaluated and changed, with the result that the last three technicians interviewed were able to identify the mathematical skill or concept represented by each of the 66 problems.

Instructor Workshops

Instructor workshops, which were part of the project unrelated to the investigation of the industry, were also used to check upon the reliability and validity of the Q-sort instrument. Each of the 33 participating instructors was asked to sort the cards in terms of the extent to which he would expect graduates of his program to master the mathematical skills and concepts of the problems. This procedure was followed by a discussion of the problems in the sort, which resulted in a further refinement of the instrument. At this point, the author felt that the Q-sort deck was sufficiently refined and reflected the mathematical continuum well enough to be used empirically in an investigation of the needed mathematical competencies in electrical-electronics industries in California.

The Ideal Sort

In addition to the sort made by the technicians in industry, which was determined by the extent to which they encountered the mathematical skills or concepts involved in the problems on the cards, there was also an ideal —or logical—sort by instructors in the area of mathematics. They sorted the cards in the logical sequence of instruction, that is, the order in which they would present the mathematical skills and concepts if they were to teach the entire range. This allowed a comparison between the sorts made by the technicians and the logical sequence of mathematics as seen by the mathematics instructors.

Interview Data

An interview sheet was designed to identify formal education and mathematics background, area of training, location of training, place and persons involved in the development of the interest in electronics, and job history. Job history was used to determine whether the technician was stable, horizontally mobile within the occupation, or upwardly mobile. In addition, the job classification was listed and operationally defined. From the operational definitions, classifications were developed which ultimately allowed grouping into nine technical occupational classifications. With this background from the interview and the Q-sort on the technicians, it was possible to identify the essential mathematics, the mathematics which is nice to know, and the mathematics of questionable value within the industry. In addition, it was possible to investigate the following:

1. The extent to which mathematics correlated with occupational mobility.

2. The extent to which formal mathematics and formal education background of the respondent correlated with the mathematics used on the job.

3. The difference in the mathematical skills needed as mentioned in the various classifications within the industry.

Statistical Analysis

Q-sort correlations differ from the customary correlations in that the entire range of responses of one individual is correlated with the entire range of responses of all individuals, rather than correlating items. The Q-sort data obtained from the technicians and instructors were correlated with one another. This computation yielded 6430 correlations with a range from —.697 to a +.968. A correlation matrix this large is difficult to interpret. To compute the mean correlation for the total group and by job classifications poses a lengthy computation. Kendall (5) has demonstrated that there is a linear relationship between the mean correlations and Kendall's W (coefficient of concordance).

The purpose of the research being explained in this example was to determine whether or not there was significant agreement among the technicians as to the rank order of mathematics from the essential to that of questionable value. Kendall's coefficient of concordance is designed to measure the extent of agreement and is applicable to Q-sort data, provided a correction is made for tied ranks.

In any given Q-sort study, the total sum of squares will be a constant for that study and, when corrected for tied ranks, will remain constant. It will vary only with the number of persons doing the sort. The over-all

mean for any Q-sort study also remains constant, which facilitates the computation of the sum of squares between items. Hence, the over-all computation of the coefficient of concordance can be done in a relatively short time; it can be computed by groups and the sums of squares between groups in the preliminary computations used in the over-all coefficient of concordance.

To conclude the example, it is sufficient to mention that it was demonstrated that all of the technicians in the sample agreed, and the agreement would have occurred by chance only 1 per cent of the time. It was further demonstrated that the ordering of mathematics by technicians agreed with the logical order of the presentation of mathematics at the .01 level.

Discussion

1. The Q-technique as a group measure has been proposed by Morsh, and Sorenson and Sheldon. The basic idea is applicable to group measure and the results of the Q-sort can be handled to reflect group opinion.

2. The Q-sort correlations can be computed easily with the use of conversion tables such as those suggested by Cohen, or by computing each correlation separately. The computations are simplified in that the mean, standard deviation, and total variance remain constant for any Q-sort study.

The resultant sort, which is arrived at by averaging the responses to each item for all of the individuals who responded, can be used as a consensus of the opinions solicited.

Kendall's coefficient of concordance can be used to demonstrate the extent to which the individuals responding agree on the order of placement of the items within the sort. An example of computations involved follows.

The basic formula of the coefficient of concordance (4) is

$$W = \frac{\text{Sum of Squares between columns}}{\text{total sum of squares}}$$

The correction for ties is

$$C = \Sigma \frac{K^3 - K}{12}$$ when K= number of items tied for a given rank in a sort.

The total sum of squares thus becomes

$$SS_{ToT} = \frac{m(n^3 - n)}{12} - C,$$ when m = number of respondents, n = number of items.

For example cited, there were 66 items forced in nine ranks of 3, 4, 7, 12, 14, 12, 7, 4, and 3 items respectively. The computation of SS_{ToT} then becomes

$$SS_{TOT} = \frac{m(66^3 - 66)}{12} - \Sigma\frac{K^3 - K}{12}$$

Since there are the same number of ties in each sort, C is equal to m times 583.5 for this example. Hence: $SS_{TOT} = 23{,}369m^2$ and

$$SS_{BC} = \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{m} Xij\right)^2}{m} - \frac{mn(n+1)^2}{4}$$

Again substituting 66 for n in the example

$$SS_{BC} = \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{m} Xij\right)^2}{m} - 74{,}068.5m \quad \text{and}$$

$$W = \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{m} Xij\right)^2}{m} - 74{,}068.5m$$
$$\overline{\qquad\qquad 23{,}369m \qquad\qquad}$$

Division results in the final equation which is applicable to all of the coefficients of concordance for the example cited.

$$W = \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{m} Xij\right)^2}{23{,}369m^2} - 3.17$$

This is applicable to groups of seven or over without significant distortion.

3. The distinct advantage of the Q-sort compared with rating scales or critical incidence methods is that the Q-technique requires that each item in the sort be compared with all other items.

4. It is feasible to include items in a Q-sort to identify those persons who are attempting to falsify their responses. This can be done by including items that are patently not applicable to the problem under consideration or by including items that are identical in concept with other items in the sort but are worded differently.

5. The Q-technique is applicable to such areas of curriculum development as the following:

a. Faculty opinion concerning the relative value of courses offered in a given curriculum.

b. Lay committee evaluation of existing curriculum content. Here it is possible to compare the lay person's opinion of

what the curriculum should include with what the curriculum actually does include.

c. The content of any set of Q-sort items can be arrived at by the participation of the persons who will ultimately respond to the Q-sort.

d. The Q-technique allows for comparison not only between individuals in terms of their opinion concerning curriculum content, but also allows for a before-and-after comparison of any given individual's opinion concerning curriculum construction due to workshops or other faculty training sessions.

BIBLIOGRAPHY

1. Allen, David. *A Guide for Developing Electronics Courses.* California State Department of Education, 1958.
2. Brandon, George L. *Twin Cities Technicians.* College of Education, Michigan State University, East Lansing, 1958.
3. Cohen, J. "An Aid in the Computation of Correlations Based on Q-sorts," *Psychological Bulletin,* 54:138-39, March 1957.
4. Edwards, Allen L. *Statistical Methods for the Behavorial Sciences.* New York: Rinehart & Company, Inc., 1954. p. 4020.
5. Kendall, M. B. *Rank Correlation Methods,* p. 81. London: Griffin, 1948.
6. Miller, Paul A. *Industrial Technicians for Greater Syracuse.* Board of Education, Syracuse 4, New York, 1960.
7. Morsh, Joseph E. "The Q-sort Technique as a Group Measure," *Educational and Psychological Measurements,* 15:4:390-95, Winter 1955.
8. Sorenson and Sheldon. "On the Use of the Q-technique in Educational Evaluation and Research," *Journal of Experimental Education,* 29:2:143-52, December 1960.
9. Stephenson, William. *The Study of Behavior.* Chicago: University of Chicago Press, 1953.

An Evaluation of Some Aspects of the Los Angeles City Schools Driver Training Program

GEORGE H. GOODY AND PHILIP G. NASH

The purpose of this study was to collect, present, and interpret material on certain aspects of the Los Angeles City Schools' driver training program which might be of help to the Board of Education in making decisions concerning the future of the program. The study was limited to driver training only—behind-the-wheel instruction conducted in an automobile. Driver education, the classroom phase of the program, was not included. The question of the efficiency of driver training instruction as now offered—how well the subject is being taught—also was outside the scope of this report, as well as was any investigation of whether successful completion of driver training actually does reduce traffic accidents or citations.

The study dealt with the following questions:

1. Does the present system of selecting pupils to take driver training result in any differences in ability and/or achievement levels between pupils who take driver training and those who do not?
2. What happens to the school grades of pupils who take driver training, and in what ways do their grades differ from those who do not take the course?
3. What do the various interested groups—school administrators, A-12 pupils, their parents, and the voters of the community—

Philip G. Nash is a research specialist in the evaluation and research section of the Los Angeles City School Districts, a position he has held for eight years. Mr. Nash, formerly a junior high school teacher in the Los Angeles city schools, received his master's degree in 1933 from the University of Southern California.

George H. Goody, also a research specialist in the Los Angeles Districts for the past two years, was formerly a high school teacher and administrator, junior college instructor and counselor, and state college assistant professor of psychology. A certified psychologist, Dr. Goody obtained his Ph.D. degree from the University of Southern California in 1951.

believe and desire with respect to the function of the high school in preparing pupils for driving?

4. To what extent are these groups willing to express themselves with respect to the financial implications of driver training?

Procedures Followed

The procedures followed in this study have been described in detail in the complete report.[1] Briefly, they were as follows:

By means of a systematic referral to test data on file in the Evaluation and Research Section and to pupil personnel records maintained by schools, a random sample of 1368 A-12 pupils from 37 high schools was obtained, of which 519 had taken driver training and 849 had not. The two groups were compared to see if there was any evidence of selectivity in the scheduling of pupils to take driver training.

The second step was an investigation of the grades earned by driver training pupils and progressive changes in their grades, if any. For this purpose, a sample of the pupils previously described was matched carefully with a similar sample of non-driver training pupils, using I.Q. as the criterion. The grade-point averages of driver training pupils before, during, and after their driver training semester were compared with the grade-point averages of the non-driver training matched controls for the corresponding periods in each subject's high school career.

The third procedure was a sampling, by questionnaire, of the experiences, attitudes, and beliefs relating to driver training of five groups of individuals:

1. A random sample of 1217 A-12's of 38 high schools, both driver training and non-driver training.

2. A random sample of 683 parents of A-12's of 37 high schools, both driver training and non-driver training.

3. A random sample of 456 voters representing the proper proportions of all areas of the Los Angeles City High School District, which included, of course, both parents and non-parents.[2]

4. The principals of 38 high schools.

5. The head counselors of 38 high schools.

---

[1]Los Angeles City School Districts. *An Evaluation of Some Aspects of the Driver Training Program in Los Angeles City High Schools.* Evaluation and Research Section Research Report No. 229. 1961.

[2]Approximately 55 per cent of the questionnaires were filled out and returned by the voters, and a little less than this amount by the parents.

Selectivity of Pupils

The results of the study of selectivity confirmed what had been suspected: pupils who took driver training were not representative of the school population. When matched with a sample of non-driver training pupils for the corresponding semesters of their high school careers, the mean I.Q. of the driver training group was 104.21, and that of the non-driver training group was 100.66. The difference between means, 3.55 I.Q. points, was found to be significant at the .01 level of confidence. This comparison was then repeated for achievement, using as the criterion standard scores on the composite of eight sub-tests of the *Iowa Tests of Educational Development*. Again the driver training group was found to be higher by a statistically significant amount.

The point here is that one should treat with caution the many published figures which show that pupils who have had driver training maintain better traffic records than do those who have not had such training. As far as pupils of the Los Angeles City schools are concerned, there has not necessarily been demonstrated any cause and effect relationship between driver training and improvement in traffic record. The improvements so often cited may, for instance, have been due to the fact that graduates of driver training courses seem to be, as has been suggested, the "conscientious, resourceful, and responsible type of individual who, perhaps, without benefit of any driver training, would still have reflected a positive effect on accident statistics."[3] However, there is no doubt that it is important to know definitely whether driver training does result in an improved traffic record. The best way to investigate this would appear to be by means of a follow-up study, in which the traffic records of equated groups could be studied over a period of two or three years following the course. Such a study is sorely needed at this time.

Effect on Grades

An additional finding of this study was that driver training has had no effect on school grades of pupils in the sample. Grade-point averages of both driver training pupils and non-driver training pupils of the sample tended to increase during this period, but such a gradual increase is not unusual. The driver-training sample of 301 pupils averaged 2.52 grade points before taking the course, and 2.65 afterward, for an average gain of 0.13 grade points. The non-driver training pupils, matched for I.Q., went from 2.46 to 2.58, for an average gain of 0.12 grade points. As a check, samples were re-matched according to grade-point average just before the

[3]Op. cit., page 3.

driver-training semester, and here again the gains in grade-point average after the course were, for driver training and non-driver training pupils respectively, .05 and .04, a non-significant difference.

To see if there were any evidences of a decline in scholarship during the driver-training semester, samples of 50 boys and 50 girls were followed for three semesters, beginning with their standing during the semester just preceding driver training. The grade-point averages for the boys were 2.42, 2.46, and 2.55. For the girls the figures were 2.61, 2.58, and 2.67. None of these variations was of sufficient amount to have significance. This finding is consistent with the expressed beliefs of principals, A-12's, and their parents, of whom fewer than one in ten thought that the driver training course had adversely affected school marks.

Car Ownership and Driving Compared

One of the findings related the prevalence of automobile ownership and driving as reported by pupils. These data were obtained from responses to a questionnaire, and their validity may be open to question; however, factors causing possible inaccuracies should be equally applicable to both groups. The data are reported in Table I.

TABLE I

Automobile Ownership and Frequency of Driving as Reported by Pupils

	Driver Training Pupils	*Non-Driver Training Pupils*
Have a driver's license	71 per cent	68 per cent
Own an automobile	31 " "	34 " "
Drive on school afternoons	57 " "	55 " "
Drive on school evenings	57 " "	58 " "

It will be apparent that there are no significant differences in any of these factors.

Opinions Reported

The remaining findings of this study represent opinions of driver training and its place in the high school as expressed by high school principals, A-12 pupils, their parents, and a sample of voters.

Of the 38 high school principals who responded, almost half wanted to continue the present driver-training program, in which the course is an elective open to qualified pupils who have completed the classroom phase, and is offered throughout the school day. Pupils are recruited for the 12 to 24 periods of the course almost entirely from study hall or physical education. An additional 18 per cent of the principals thought that the driver-training program should be expanded, so that it may be said that two-thirds of the principals were favorably disposed toward the program. About one-third were dissatisfied with the program as it then stood, but most of these would accept it if "substantial changes" were made.

Sixty-one per cent of the principals believed that, if a given amount of money for a school's budget were specified, expenditures for driver training should remain just about as they are now. One principal believed that expenditures for driver training should be increased, even if it necessitated a decrease for some other phase of the educational program. Thirty-six per cent thought expenditures for driver training should be decreased, and those for some other phases of the educational program increased. Opinions of these principals appeared to be more varied than were those of other groups, with their concern centering around, first, the relative expense, and secondly, the teacher and pupil time which driver training took from the regular school day.

Opinions of other respondents were more favorable toward driver training. Of the 550 A-12 pupils in the sample who had had driver training, 96 per cent considered it "a valuable experience." Of the 667 pupils who had not had the course, 70 per cent expressed themselves as wishing they had had an opportunity to take it. Ninety per cent of the parents of A-12's and 88 per cent of the voters, including both parents and non-parents, believed that it is the "business of the high school" to furnish driver training instruction.

The high per-pupil-hour cost of driver training instruction has been, of course, of great concern. To be sure there would be no misunderstanding, the item on cost in the questionnaires was presented as follows:

Taking into consideration the teacher's salary only, it costs the school on the average 19 cents per pupil per hour to teach regular classes, but $1.98 per pupil per hour to teach behind-the-wheel driver training classes. In view of this fact, do you believe that driver training should be taught by the high school?

Eighty-five per cent of the parents answered "yes" to this item, as did 72 per cent of the sample of voters. The voter sample had been identified in the questionnaire as "parent" or "non-parent" of a pupil in a Los Angeles City school, and affirmative response was noted from 86 per cent of the parent-voters and from 65 per cent of the non-parents in the voter sample.

It is noteworthy that there was no noticeable drop in the essentially

favorable opinion of the respondents when the cost factor was presented. Virtually the same high proportions of parents and of voters believed that driver training should be taught and should be supported financially in the public schools of Los Angeles regardless of cost.

A variety of reasons for maintaining driver training classes were advanced. It was evident from responses that most people believe sincerely that someone must act to improve the traffic situation, and that they have faith in the high schools to do it. Comments such as these were commonly found:

A-12 pupil: **"The program is excellent; everyone who drives should take it."**

Parent: **"The cost of blood spilled on our highways is far more than this behind-the-wheel instruction cost, no matter how much it might be."**

Voter: **"I think driver training in the high schools is one of the best things that could ever happen to our young people. Things they learn properly at that age stay with them always."**

Voter: **"This cost will be placed on the individual's shoulders in one way or another. Let's put it in the classroom where qualified teachers can control it."**

Also high on the list of reasons of both pupils and parents was the desire to participate in the lowered insurance premium rate which is offered to graduates of the driver-training course. This financial consideration is probably the most compelling one for pupils and parents.

Most respondents believed that the most unsatisfactory element of the present driver-training program in Los Angeles City high schools is the fact that many who wished to take it were unable to be accommodated. At present there is no uniform system for selecting driver-training applicants, and each high school sets up its criteria. These procedures are obviously not well understood by either pupils or parents. Unless the present format for conducting driver training is changed so that all qualified applicants can be accepted, early consideration might well be given to a uniform method of selecting pupils to take the course so that the fairness of selection procedures is evident to all.

Research News and Views

Garford G. Gordon, chairman of the California Advisory Council on Educational Research, temporarily has left his position of Research Executive of the California Teachers Association to direct a special project in Karachi, Pakistan, under the auspices of UNESCO. Dr. Gordon was granted a leave of absence from the Association after his selection for this post, and will be abroad for at least a year. Although there has not been time for Dr. Gordon to send word regarding the details of his assignment, it is expected that his duties will involve development of procedures for the collection and compilation of data on various aspects of Pakistan's educational system. Before departing for Karachi in mid-August, Dr. Gordon worked with the UNESCO in Paris for two months on a World Education project.

During Dr. Gordon's absence, the CTA's research department will be headed by Dr. John H. Bright, Assistant Research Executive, who will also act as temporary chairman of CACER

Education Index will commence listing the contents of the *California Journal of Educational Research* with this issue, after election of the publication by the *Index* subscribers. *Education Index* is received by approximately 2700 subscribers in the United States and abroad and has the reputation of being one of the most valuable tools in the field of education and educational research. The *CJER* editors are grateful to the *Index* subscribers who expressed their favorable opinion of this publication.

Julius H. Stier, Director of Research for the Santa Monica Unified School District, has just been added to the membership roster of the California Advisory Council on Educational Research, to fill the vacancy created by the retirement of Dr. Henry Weitzel from the Pasadena system. Dr. Stier has filled his present position for the past 15 years and has had, as well, experience in all levels of education. He has been an elementary teacher, high school teacher, high school district superintendent, senior vocational counselor at University of California, Los Angeles, and lecturer at both University of California Extension and University of Southern California. Dr. Stier obtained his Ed.D. degree in 1956 from the University of Southern California. He is currently Secretary-Treasurer of the California Educational Research Association, and was General Chairman in 1960 of the Southern Regional Guidance Conference.

Everett B. Chaffee has been designated by the California Association of School Administrators to be its official representative on the California

Advisory Council on Educational Research. The Executive Secretary of CASA, Dr. James Corson, will participate in CACER activities as his duties permit, but the selection of Dr. Chaffee will enable the administrative organization to work more actively and closely with the Advisory Council. For the past five years Associate Superintendent, Division of Instructional Services, Los Angeles City Schools, Dr. Chaffee's former experience includes that of teacher, vice-principal, principal and assistant superintendent with the Los Angeles Districts. Dr. Chaffee has been an active participant in other professional associations, which include the California Educational Study Council, the California Teachers Association, the California Association of Secondary School Administrators, and the University of Redlands Alumni Association.

Research for Educational Planning is the theme of the Thirteenth Annual State Conference on Educational Research, to be held at the Ambassador Hotel in Los Angeles on November 3 and 4. The program includes a clinic on Use of Computers, preceded by guided tours of the IBM Computer facilities. The two-day program will contain twelve section meetings designed to cover all aspects of the educational program. Information and pre-registration forms are available from the CACER Conference Office, 1705 Murchison Drive, Burlingame, California.

The CACER Conference will be preceded by a one-day meeting of the new California Educational Data Processing Association. Information regarding the Association and the November 2 meeting can be obtained from Dr. Alvin Grossman, State Department of Education, 721 Capitol Avenue, Sacramento 14, California.

Philosophy and Research

(Continued from page 146)

a small college in the Middle West. The professor, in thumbing through a primer brought home by his son, found three words he had never met before. He thereupon interviewed the other doctors of philosophy among his colleagues and found that most of them did not know the words. In writing his report, the professor thought it would be carrying coals to Newcastle to point out the implications of his findings. He saw no possible meaning to his findings other than that the three words were inappropriate for primary readers. However, one of the other doctors said, not facetiously, that the findings proved that primary pupils of one generation were more broadly cultured than doctors of philosophy of the preceding one.

With research's dependence on philosophical guidance, it often might be well for researchers to remark near the ends of their reports somewhat as follows: **"In the light of the underlying assumptions and criteria of this investigation, the conclusions seem to be . . ."**